

贪心与回溯算法在城市马拉松路线规划中的实践^①



王友才^{1,2}, 陈焱焱¹, 徐玉兵¹, 刘子含^{1,2}, 潘瑞^{1,2}, 何子军¹, 杨先军¹, 孙怡宁¹

¹(中国科学院合肥物质科学研究院, 合肥 230031)

²(中国科学技术大学, 合肥 230026)

通信作者: 陈焱焱, E-mail: yichen@iim.ac.cn

摘要: 针对目前城市马拉松路线人工规划效率低下的问题, 本文采用贪心和回溯算法进行城市马拉松路线智能规划, 具体方法是: 通过城市路网信息构建由经纬度坐标点拓扑关系连接而成的路网, 采用贪心和回溯算法对坐标点进行遍历搜索, 结合城市马拉松路线特殊要求, 运用直接逼近、启发式远离、启发式靠近和方向预估等策略实现路线的智能规划. 在此基础上, 提出一种综合 POI 热度值、道路宽度适宜度、路线畅通指数、过弯舒适度以及 POI 密集度 5 个维度的马拉松路线评估方法. 最后, 开展了北京、合肥马拉松人工和智能规划路线对比分析, 结果表明所采用的方法可快速高效实现马拉松路线规划.

关键词: 城市马拉松; 路线规划; 评估体系; 启发式策略; 贪心算法; 回溯

引用格式: 王友才, 陈焱焱, 徐玉兵, 刘子含, 潘瑞, 何子军, 杨先军, 孙怡宁. 贪心与回溯算法在城市马拉松路线规划中的实践. 计算机系统应用, 2023, 32(3): 275-281. <http://www.c-s-a.org.cn/1003-3254/8990.html>

Practice of Greedy and Backtracking Algorithm in City Marathon Route Planning

WANG You-Cai^{1,2}, CHEN Yan-Yan¹, XU Yu-Bing¹, LIU Zi-Han^{1,2}, PAN Rui^{1,2}, HE Zi-Jun¹, YANG Xian-Jun¹, SUN Yi-Ning¹

¹(Hefei Institutes of Physical Science, Chinese Academy of Sciences, Hefei 230031, China)

²(University of Science and Technology of China, Hefei 230026, China)

Abstract: Manual planning of city marathon routes has low efficiency. In view of this, this study adopts a greedy and backtracking algorithm to carry out intelligent planning of a city marathon route. The specific method is described as follows. A road network connected by the topological relationship of longitude and latitude coordinate points is built through the urban road network information, and a traversal search is performed by the greedy and backtracking algorithm on the coordinate points. In addition, according to the special requirements of the city marathon route, strategies are adopted, such as direct approximation, heuristic distance, heuristic approach, and direction estimation, so as to realize the intelligent planning of the route. On this basis, a marathon route evaluation method is proposed, which integrates five dimensions including POI heat value, road width suitability, route smoothness index, comfort for turning, and POI density. Finally, a comparative analysis of artificial and intelligent route planning for Beijing and Hefei marathons is carried out. The results show that the proposed method can realize fast and efficient marathon route planning.

Key words: city marathon; route planning; evaluation system; heuristic strategy; greedy algorithm; backtracking

近年来, 中国马拉松赛事发展迅速, 在中国田径协会注册的马拉松赛事经历了从 2010 年 13 场到 2019

年 1 828 场的跨越式发展, 全国马拉松赛事参加人次从 2013 年的 75 万增长到 2019 年的 712 万, 马拉松已

① 基金项目: 国家重点研发计划 (2021YFF0306402); 安徽省科技重大专项 (202103a07020004)

收稿时间: 2022-07-23; 修改时间: 2022-08-26; 采用时间: 2022-09-27; csa 在线出版时间: 2022-11-29

CNKI 网络首发时间: 2022-11-30

成为国内大众参与的热门赛事^[1]。马拉松赛事的举办对凝练城市文化、提高城市软实力、城市经济发展以及民生问题的提升都有着积极的促进作用^[2]。马拉松赛道路线一般由主办方根据地形图以及城市特色文化地标进行手动规划路线,最后利用装有琼斯计数器的校准自行车进行赛道测量^[3]。马拉松路线规划不仅需要耗费大量的资源,而且人工规划的局限性也会导致规划的路线存在诸多安全隐患,例如赛道拥挤、交通拥堵等^[4,5]。

对于路线规划的问题,常用的算法有贪心算法、回溯算法、Dijkstra 算法、A*算法、蚁群算法等,这类算法解决的是最优化问题,例如求最短路径、最短时间等。马拉松路线规划要满足距离为 42.195 公里,且起终点之间的直线距离不得超过总距离的 50% 等要求^[6]。针对马拉松路线规划,黄鹤等提出改进的 A*算法对马拉松路线进行规划^[3],有效地简化了马拉松选线工作中的繁琐步骤,使马拉松路线的选线更加高效、直观。但 A*算法是在两点间生成一条最短路径,无法得到往返路径而实际路网需要考虑道路双向问题,因而存在一定局限性。

本文结合城市路网、POI 数据以及城市道路信息对城市马拉松路线规划进行研究,以经纬度坐标点反应路网信息,以道路交叉口作为路线转移节点,可以充分考虑道路的双向问题和转弯角度情况,减小算法生成路径时的计算成本。同时,基于贪心和回溯算法,在每个路口处,根据启发式预估策略获得路线局部最优解从而得到最终备选路线。在此基础上,提出一套评估体系对赛道路线进行评估。最后通过规划路线与人工路线对比分析验证方法的有效性。

1 总体思路

城市马拉松路线规划必定满足以下必要条件: 1) 从起点出发经过不重复的路段最终到达选定终点; 2) 路线总距离为 42.195 公里。但是在城市路网中从起点到终点之间存在诸多拓扑关系,所以该问题的解空间庞大且无最优解。为更快得出路线,给出包括起终点在内的几个关键点,将该问题近似的变成旅行商问题。即使如此,依旧存在 42.195 公里的限制,且关键点与关键点之间不是最短路径的优化问题。为了得到一个符合的路线,在城市路网中利用贪心思想和回溯剪枝的手段得到局部最优解,从而得到全局解^[7]。如图 1 所示,整个方法分为 3 个部分: 构建路网、规划路线和评估路线,其中规划路线包括 4 个核心模块: 转弯角度测量及相对位置定位、启发式预估策略、方向预估策略和靠右原则。

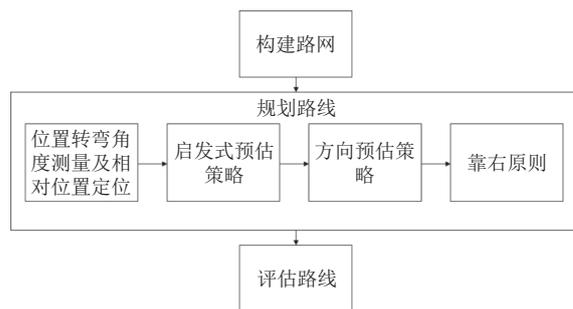


图 1 方法流程图

通过分析城市路网信息可知整个路网由一个个经纬度坐标点根据拓扑关系连接而成,每一条道路由多个路段组成,每个路段由多个坐标点构成,根据每个坐标点的连接情况,将坐标点划分为单联节点、双联节点和多联节点,如图 2 所示。本文提出路线规划方案是根据坐标点遍历搜索法进行,单联节点作为路线搜索回退标志,双联节点作为路线折返标志,多联节点作为路线转移标志和路线折返标志。

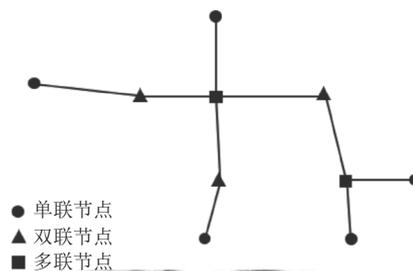


图 2 路网节点拓扑关系及节点分类

针对马拉松路线的特殊性,需要满足一定的起跑距离以达到更好的赛事效果,因此将搜索分为两个阶段: 起跑阶段和非起跑阶段^[8]。规定从起点出发,在起跑阶段选择直道路线并持续一定距离,之后到达非起跑阶段。如图 3 所示,在非起跑阶段,在当前路口节点处进行判断决定采用何种策略进行下一个坐标点的查找。本方法给出 3 种启发式预估策略: 直接逼近策略、启发式远离策略和启发式靠近策略^[9]。

这 3 种策略每一个策略都要满足如下条件: 1) 为保证规划舒适的赛道路线,要确保下一个点转弯角度大于 75° ^[8]; 2) 为避免路线交叉,要确保不走已走过的坐标点; 3) 为避免搜索提前结束,要确保不提前进入终点所在路段。满足以上条件则更新状态,否则进行回溯,回到上一个状态,策略检查流程如图 4 所示。

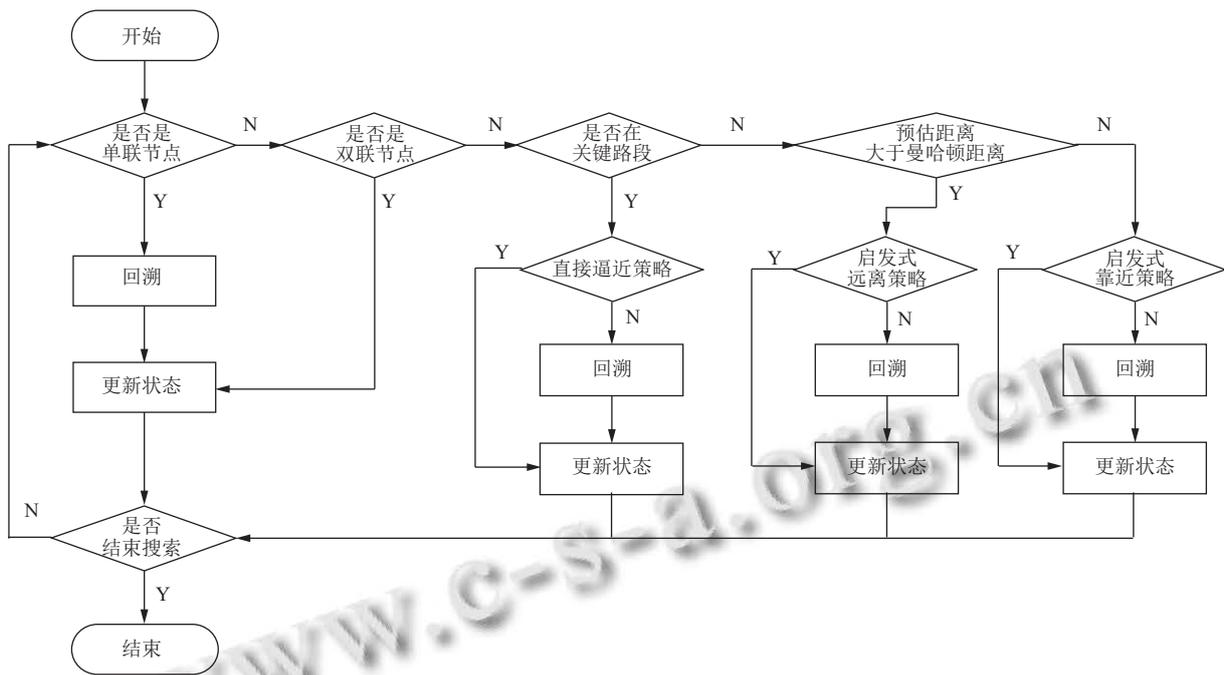


图3 非起跑阶段的路线搜索流程图

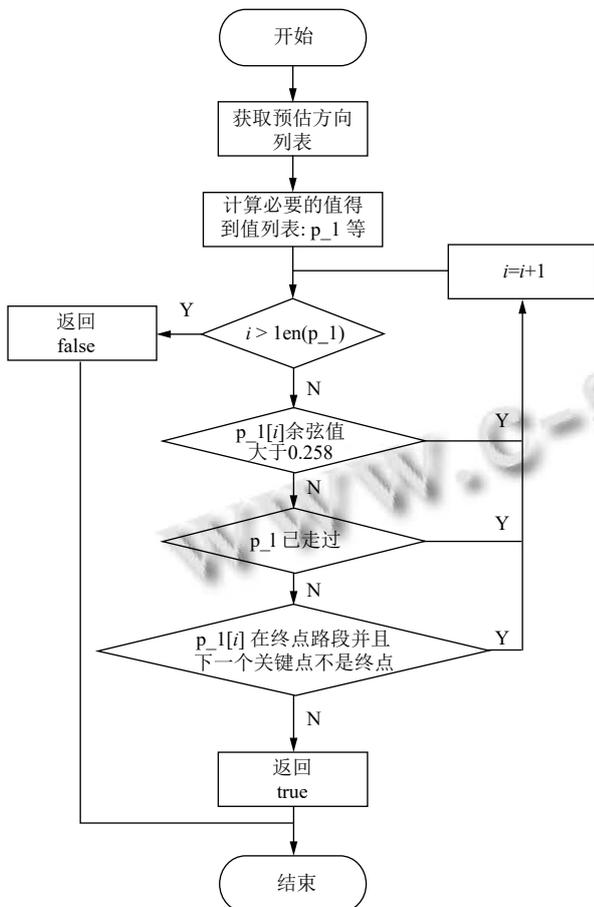


图4 策略检查流程图

2 路线规划

2.1 转弯角度测量及相对位置定位

本文中的路网是由一个个经纬度坐标点连接而成, 所以可以针对连续的3个坐标点, 通过计算其余弦值来反映转弯角度 (假设3个坐标点在同一平面内), 这有助于在道路交叉口进行路线转移.

$$\cos = \frac{(x1 - x2)(x3 - x2) + (y1 - y2)(y3 - y2)}{\sqrt{(x1 - x2)^2 + (y1 - y2)^2} \sqrt{(x3 - x2)^2 + (y3 - y2)^2}} \quad (1)$$

其中, \cos 表示从坐标点 $(x1, y1)$ 经坐标点 $(x2, y2)$ 到坐标点 $(x3, y3)$ 所形成的转弯角度的余弦值.

路网中的道路不都是横平竖直, 因此在规定路线方向时, 约束一定角度范围为某一固定方向, 通过计算两坐标点 (假设两个坐标点在同一平面内) 的斜率来反映两点形成的直线相对于横坐标轴的倾斜程度, 从而确定它们的相对位置, 这有助于判断在路口处是否右转. 相对位置以上、下、左、右、左上、左下、右下和右上表示.

$$\tan = \frac{y1 - y}{x1 - x}, x1 \neq x \quad (2)$$

其中, \tan 表示坐标点 $(x1, y1)$ 与坐标点 (x, y) 的正切值, 即斜率.

2.2 启发式预估策略

直接逼近策略是当前坐标点在下一个关键点路段

时为避免错过关键点而直接逼近的策略,其预估方向为下一个关键点所在的方向;启发式远离策略是在到下一个关键点前预估距离大于曼哈顿距离时而采取的策略,此时为保证全程的距离接近 42.195 km 而选择远离关键点的方向;启发式靠近策略是在到下一个关键点前预估距离小于等于曼哈顿距离时而采取的策略,曼哈顿距离近似代表当前距离到达关键点最短距离.在二维平面两点 $a(x_1, y_1)$ 与 $b(x_2, y_2)$ 的曼哈顿距离计算公式如下:

$$M(a, b) = |x_1 - x_2| + |y_1 - y_2| \quad (3)$$

为实现有效的启发式靠近和远离策略,需要动态的计算当前坐标点到下一个关键点剩余的曼哈顿距离和剩余预估距离.路网中关键点拓扑示意图如图 5 所示.

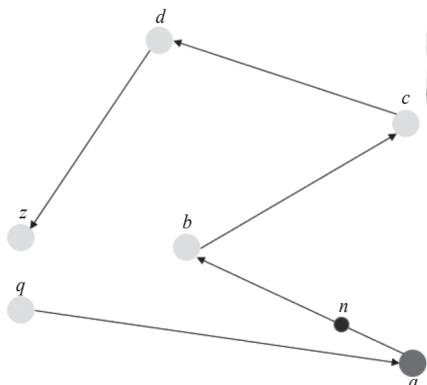


图 5 关键点拓扑示意图

图 5 中, q 为起点节点, z 为终点节点, a, b, c, d 为关键点节点, n 为当前节点. 当前点到下一个关键点剩余的曼哈顿距离计算公式如下:

$$RestD = \frac{D_{sum} - \sum_{i=0}^{j-1} D(K_i, K_{i+1}) - \sum_{i=j}^{n-1} M(K_i, K_{i+1})}{\sum_{i=j}^{n-1} M(K_i, K_{i+1})} \times M(K_j, K_{j+1}) + M(K_j, K_{j+1}) - D(K_j, K_{j+1}) \quad (4)$$

其中, n 表示关键点的个数, j 为当前关键点的位置, K_i 为第 i 个关键点, D_{sum} 表示总的路线距离, $M(K_i, K_{i+1})$ 表示第 i 个关键点至第 $i+1$ 个关键点的曼哈顿距离, $D(K_i, K_{i+1})$ 表示第 i 个关键点至第 $i+1$ 个关键点的已走距离, $\sum_{i=0}^{j-1} D(K_i, K_{i+1})$ 表示从起点到当前关键点的已走距离, $\sum_{i=j}^{n-1} D(K_i, K_{i+1})$ 表示当前关键点至终点的曼哈顿距离. 当前坐标点至下一个关键点要走的剩余曼哈顿距离计算公式如下:

$$RestM = M(p, K_{j+1}) \quad (5)$$

其中, p 为当前坐标点.

2.3 方向预估策略

通过启发式预估策略已经知道在路口处接下来是靠近还是远离下一个关键点,但是到达一个路口时有多个方向可选,对于方向的选择提出了方向预估策略,根据当前坐标点与关键点的位置以及当前方向给出预估方向列表,并按照优先级排序,以确保局部最优.对靠近和远离两种情况提出靠近原则和远离原则.靠近原则优先推荐与关键点相同的方向,其次推荐靠近关键点且与当前方向相同的方向,最后推荐靠近关键点但与当前方向不同的方向;远离原则优先推荐与关键点相反的方向,其次推荐远离关键点且与当前方向相同的方向,最后推荐远离关键点但与当前方向不同的方向.

如图 6 所示,当前方向向下,关键点位于当前点的左下方,接下来有左,右以及下 3 个方向可走.根据方向预估策略,若靠近关键点,则预估方向列表为:左下、下、左、左、下;若远离关键点,则预估方向列表为右上、上、右、上.

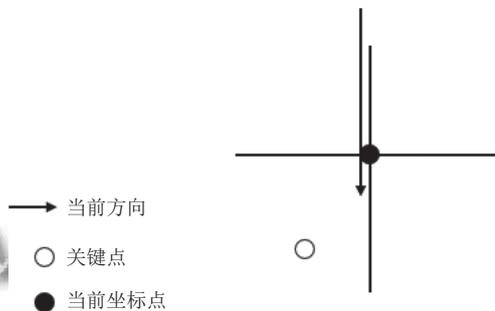


图 6 坐标相对位置示意图

2.4 靠右原则

通过方向预估策略已经得出路线转移方向.由于本方法充分考虑路网中道路双向的问题,为避免路线搜索混乱,本文约束对于一条道路选择前进右侧的路段,即靠右原则.

路口示意图如图 7 所示,路口的具体控制策略流程图如图 8 所示,采取的策略是通过当前方向找到下一个与当前方向相同的点并判断其是否是路口节点,如果是路口节点则通过该节点进行转弯决策.若决策是右转则不过该路口节点,若决策是左转,则穿过该路口节点.

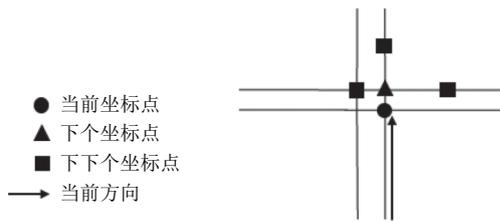


图7 路口坐标点示意图

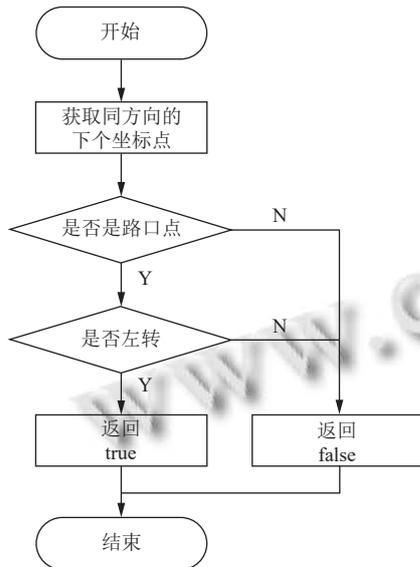


图8 路口左转判断流程图

3 评估体系

城市马拉松赛道是宣传一个城市文化、科技和经济的重要媒介^[10], 同时作为赛事的核心, 赛道要满足赛事要求, 为选手提供舒适安全的比赛环境, 也需降低对城市交通的影响. 根据国内外学者对马拉松赛道的分析和权威机构发布的文件^[11], 本文对每一条路线从 POI 热度值、道路宽度适宜度、路线畅通指数、过弯舒适度以及 POI 密集度 5 个维度进行综合评估, 其计算公式规定如下:

$$S = \frac{(S_{poi} + S_{road} + S_{cross} + S_{cos} + S_{poin})}{5} \quad (6)$$

其中, S 是路线的最终评估分数, S_{poi} 是路线 POI 热度值得分, S_{road} 是道路宽度适宜度得分, S_{cross} 是路线畅通指数得分, S_{cos} 是过弯舒适度得分, S_{poin} 是 POI 密集度得分.

城市 POI 数据含有城市的地理、地貌以及人文特征, 能够凸显城市形象和赛道文化^[12]. 定义 POI 热度值为赛道路线附近的所有 POI 数据的热度值均值. 将每个 POI 数据定义初始值为 100 分, 先按照大类进行等

次划分, 再按照小类进行等级划分, 进一步量化 POI 热度值, S_{poi} 计算公式如下:

$$S_{poi} = \frac{\left(\sum_{i=1}^{n_p} 100 \times G_i \times L_i \right)}{n_p} \quad (7)$$

其中, G_i 表示第 i 个 POI 数据的等次所持比例, L_i 表示第 i 个 POI 数据等级所持比例, n_p 是路线所包含的 POI 数据数量.

关注 POI 质量的同时也关注 POI 数量. 本文对近几年国内著名马拉松路线研究^[13], 分析其赛道 POI 数量, 总结 POI 密集度得分 S_{poin} 计算公式如下:

$$S_{poin} = \begin{cases} 0, & n_p < 5 \\ 30, & 5 \leq n_p < 10 \\ 60, & 10 \leq n_p < 20 \\ 80, & 20 \leq n_p < 30 \\ 90, & 30 \leq n_p < 50 \\ 100, & n_p \geq 50 \end{cases} \quad (8)$$

马拉松赛道宽是影响选手成绩和参赛体验的重要因素^[14]. 一条路线由多个路段组成, 规定每条路线的宽度初始值为 100, 给不同宽度的路段分配不同的占比来量化宽度适宜度. 整个路线的宽度适宜度 S_{road} 计算公式如下:

$$S_{road} = 100 \times \left(\frac{D_{\geq 12}}{D_{sum}} \times 100\% + \frac{D_{\geq 9 \& < 12}}{D_{sum}} \times 70\% + \frac{D_{< 9}}{D_{sum}} \times 50\% \right) \quad (9)$$

其中, $D_{\geq 12}$ 为路线中宽度大于等于 12 的路段总长度, $D_{\geq 9 \& < 12}$ 为宽度在 [9, 12) 区间的路段总长度, $D_{< 9}$ 为宽度小于 6 的路段总长度.

进行马拉松赛事举办时, 需要进行交通管制, 对赛道进行封路, 这势必会对公众出行造成一定的影响^[15]. 路线畅通指数反应马拉松赛道给城市交通带来的影响. 规定每条路线的路线畅通指数初始值为 100, 给不同拥堵程度的路段分配不同的占比来量化路线畅通指数. 整个路线的路线畅通指数 S_{cross} 计算公式如下:

$$S_{cross} = 100 \times \left(\frac{DC_4}{D_{sum}} \times 10\% + \frac{DC_3}{D_{sum}} \times 50\% + \frac{DC_2}{D_{sum}} \times 70\% + \frac{DC_1}{D_{sum}} \times 100\% \right) \quad (10)$$

其中, DC_4 是路线中拥堵级别为“严重拥堵”的路段总长度, DC_3 是拥堵级别为“拥堵”的路段总长度, DC_2 为拥堵级别为“缓行”的路段总长度, DC_1 是拥堵级别为“畅通”的路段总长度, D_{sum} 是路线所有路段总长度 (道路的拥堵等级根据高德地图开放平台划分).

马拉松赛道中弯道路线是影响选手成绩的因素^[16],因此把弯道舒适度作为评价指标之一,规定每条路线的弯道舒适度初始值为100,给不同弧度的弯道分配不同的占比来量化弯道舒适度,整个路线的弯道舒适度得分 S_{cos} 计算公式如下:

$$S_{cos} = 100 \times \left(\frac{N_{>0}}{N_{sum}} \times 0\% + \frac{N_{>-0.5 \& \leq 0}}{N_{sum}} \times 50\% + \frac{N_{>-0.93 \& \leq -0.5}}{N_{sum}} \times 70\% + \frac{N_{\leq -0.93}}{N_{sum}} \times 100\% \right) \quad (11)$$

其中, $N_{>0}$ 是路线中过弯角度小于 90° 的弯道数量, $N_{>-0.5 \& \leq 0}$ 是过弯角度大于等于 90° 且小于 120° 的弯道数量, $N_{>-0.93 \& \leq -0.5}$ 是过弯角度大于等于 120° 且小于

150° 的弯道数量, $N_{\leq -0.93}$ 是过弯角度大于等于 150° 的弯道数量, N_{sum} 是路线中弯道总数量。

4 实验验证

在 OpenStreetMap 平台爬取了合肥市城市路网信息,在合肥市路网中,以嘉陵江路与庐州大道路口为起终点,依次以环湖北路,环湖北路与万年埠路路口,中山路与包河大道路口,庐州大道与扬子江路路口以及中山路为关键点进行路线搜索,得到马拉松路线规划路线,对比同样以嘉陵江路与庐州大道路口为起终点的2021合肥马拉松路线,如图9所示。通过对两张路线图评估和分析,得到结果,如表1所示。



(a) 规划的合肥马拉松路线图



(b) 2021 合肥马拉松路线图

图9 合肥马拉松路线图

表1 以嘉陵江路与庐州大道路口为起终点的路线评估对比表

路线	POI 密集度	POI 热度值	宽度 舒适度	路线畅通 指数	弯道 舒适度	总分
规划路线	80.00	84.62	74.40	100.00	96.75	87.15
已有路线	90.00	78.51	70.39	97.11	95.35	86.27

两个路线相似度为92.94%,所以两个路线的分值接近,本方法规划的路线包含更长的庐州大道,因此在宽度舒适度指标要多高于已有路线。

还以同样的步骤对北京路网进行处理,并选择天安门广场东侧路作为起点路段,奥林匹克公园的景观大道作为终点路段,利用本方法得到马拉松路线规划

路线,对比同样以天安门广场东侧路作为起点路段、奥林匹克公园的景观大道作为终点路段的2017北京马拉松路线,如图10所示。通过对两张路线图评估和分析,得到结果,如表2所示。

两个路线相似度为97.04%,其分值也非常相近,二者主要差别在于弯道舒适度指标,在路网中可以看到,在快接近终点处,规划路线的弯道角度更小,因此分值低于已有路线。

通过两组对比分析,表明用本研究可有效地规划出马拉松赛道路线,并且规划路线与已有路线进行对比有较大的相似之处,然而由于方法的程序化也导致与人工规划的路线存在个别差异。

